Defence Science Journal, Vol. 58, No. 4, July 2008, pp. 565-572

Auto-video Tracking System: Performance Evaluation

B.S. Chauhan, Manvendra Singh, V.K. Sharma, and P.C. Pandey

Instruments Research & Development Establishment, Dehradun-248 008

ABSTRACT

Automatic target tracking systems are employed in a wide variety of missions and tracking environment such as fire control, guidance, navigation, passive range estimation, and automatic target discrimination. The tracker performance depends upon target size, contrast, speed, and signal-to-noise ratio. The evaluation of a tracker system involves lengthy field trials and measurements. In the present article, a method for quick evaluation of tracker system and working out selection criteria for different tracking algorithm for various target and background combinations have been suggested. Performance measures such as aiming point error, duration of successful tracking, number of tracking losses, indication of confidence, and system reaction time have been used to evaluate the performance of a tracking system.

Keywords: Automatic target tracking, auto-video tracking system, tracker system evaluation, target-tracking algorithms

1. INTRODUCTION

The purpose of an auto-video tracking system (AVTS) is to maintain a stable sensor-to-target line of sight (LOS) in the presence of relative target motion and base motion disturbance to the sensor platform. The operator acquires the target, within the track gate, using joystick. After the target is acquired, the tracking sub-system locks onto it and thereafter maintains the LOS automatically.

A wide range of target characteristics can be anticipated in the typical scenario of a low-contrast target in a complex scene. Various target-tracking algorithms such as edge, centroid and correlation are available for generating the error signals wrt centre of field-of-view (FOV)[1-3]. Tracker design is environmental-sensitive. Different types of tracking systems are designed to meet the respective tracking environment. For example, the star tracker would not be able to track the manoeuvering, non-cooperative target. Similarly, an up-looking tracker designed to track airborne/spaceborne target against a sky background would not be able to track target against ground clutter in down-looking surveillance system. The performance of tracking system also depends on operator's skills. It is therefore essential to examine different target and background conditions for evaluating the AVT[4-6] system.

The requirement of target size, contrast, update rate for tracking various ground and aerial targets under different atmospheric conditions has been worked out. Various target evaluation parameters are defined and method for selection of tracking algorithm for a given scenario has been worked out.

2. TRACKING SYSTEM

Figure 1 shows the AVTS block diagram. A video tracker receives the video information from

Received 19 January 2007, revised 21 November 2007

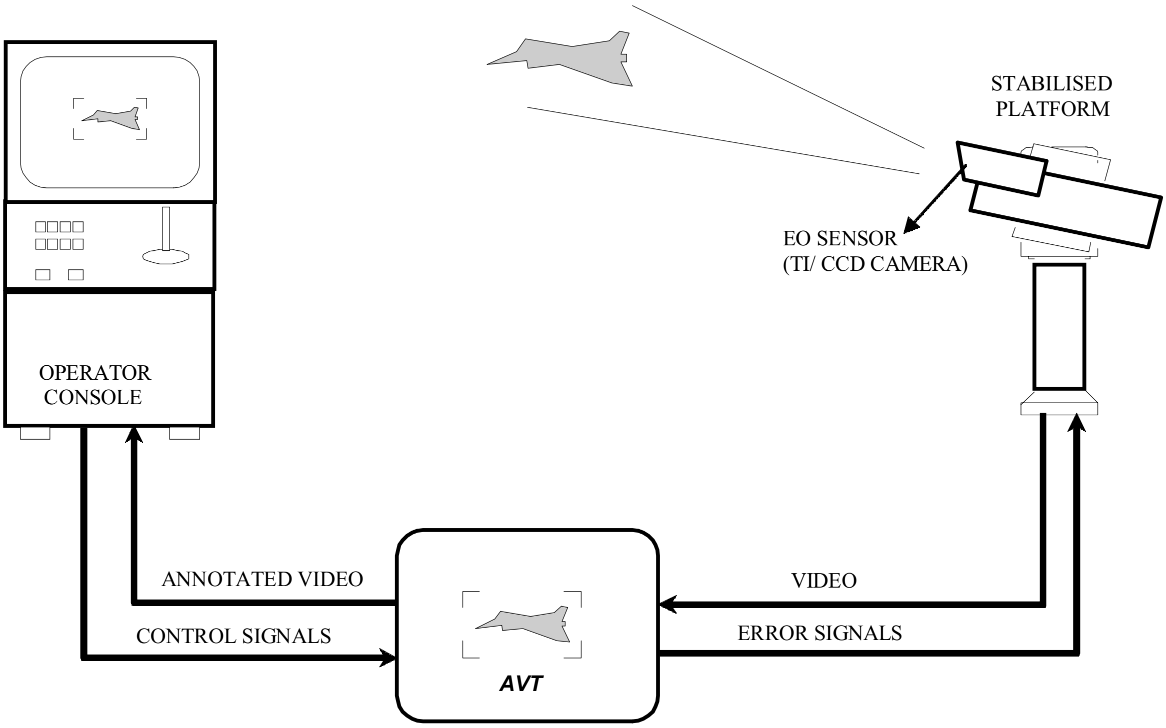

Figure 1. System block diagram of automatic video tracker.

a camera and locks it on a selected target. A feedback control loop called the track loop continuously adjusts the sensor platform to keep the target in the centre of the sensor FOV or track gate. The typical specification of a video tracker is shown in Table 1.

3. TARGET AND BACKGROUND PARAMETER

The scene parameters, i.e., target size, contrast, speed, and SNR determine the tracker performance. For tracking highly manoeuvering aerial target, update rate is also an important parameter. Effect of various target parameters on tracking performance and auto thresholding techniques for separating target from background are discussed.

Table 1. Specification of a typical auto-video tracker system

Parameter	Value
Video input	: 1 V_{p-p}, 625 lines, 50 field/s
Target to bore sight error update rate	: 50 Hz
Minimum target contrast	: 5 %
Minimum signal-to-noise ratio	: 4
RMS noise on target position output	: < 0.5 pixel
Minimum target size	: 6 × 3 pixels
Maximum target size	: 80 % of FOV
Tracking algorithm	: Centroid/correlation/edge
Tracking rate	: ± 4 pixel/ field
Output (error signal)	: 8 bit digital data (both in X and Y axis)

3.1 Target Size

Target acquisition capability of an AVT system at different ranges is dependent on target size. As per Johnson's criteria 1/4/6 cycles across target are required for 50 per cent probability of detection/ recognition/identification[7,8]. The range versus target size, generally defined in terms of target television lines (TVL) across target per field are given by:

$$\text{TVL across target/field} = \frac{\left(\begin{array}{c}\text{(target size/range)} \times 57.3 \times \\ \text{(resolution of imaging system in TVL/2)}\end{array}\right)}{\text{FOV}(^\circ)}$$

This equation is used for system design. The angle subtended by the target at a given range is calculated. Depending upon the requirement of detection/ recognition/identification of a given size target (at a given range), this angle is equated to 1, 4, 6 time of IFOV and required sensor IFOV is calculated. Alternately for given sensor parameters, one can determine whether the target can be detected or tracked at a specified range. This equation can be used for evaluation of detection and tracking performance of an imaging system with following manner:

- If the above equation results in TVL/field <1, the target cannot be detected by the imaging systems.
- If the above equation results in TVL/field =1, the target can theoretically be detected by the imaging systems.
- If the above equation results in $2 \le$ TVL/field ≤ 3, the target can be tracked but represents generally poor response from the sensor.
- If the above equation results in TVL/field ≥ 4, the target can be easily tracked and sensor performance is nearly 100 per cent.

Using Eqn (1), the range performance of an imaging system with horizontal field of view (HFOV) 3.8° and resolution 500 TVL is calculated. TVL across the target (TVL/field) as a function of range for different ground and aerial target is shown in Fig. 2. From Fig. 2 it can be seen that aircraft target at range 5 km and petrol boat at 16 km range will have 4 TVL/field. This shows that imaging system with above parameters, can track reliably aerial targets such as aircraft upto 5 km whereas the large target such as petrol boat (10 m x 10 m) can be tracked up to 16 km.

3.2 Scene Contrast

In addition to target size, the contrast of the target is an important consideration. The contrast between target and background is defined as

$$C = \frac{\alpha(M_t - M_b)}{(M_t + M_b)} + \frac{\beta(\sigma_t - \sigma_b)}{(\sigma_t + \sigma_b)} \qquad (2)$$

where M is the mean value of the intensity variation of target (t) and background (b) and σ is the measure of intensity variation of target (t) and background (b).

This value is calculated for each frame with the reference point position taken at the centre for the target and background area. α, β are scene-dependent factors[10]. The values of weighting factors have atmospheric and visibility dependency and the values ranges from 0 to 1 with $\alpha + \beta = 1$. For uniform scene, α should have typical value of 0.7, and for large variation in intensity value within the scene β should have large value. Centroid tracking separates target from the background on the basis of its contrast from background and it require minimum 5 per cent contrast therefore small targets (e.g., helicopter at a distance) need special processing.

3.3 Signal-to-Noise Ratio

The minimum target contrast can best be explained in terms of signal-to-noise (S/N) ratio for electro-

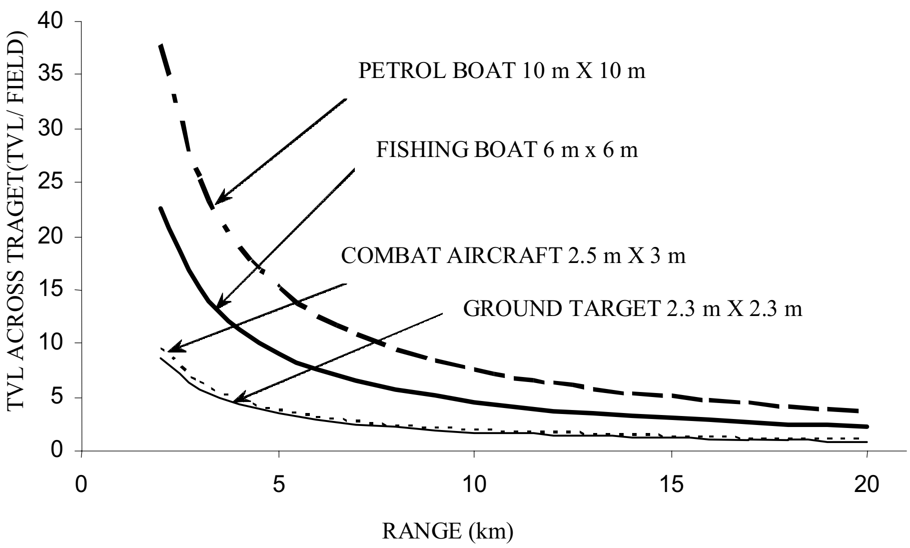

Figure 2. TVL across target at different ranges.

optical sensor system. The SNR (peak-signal-to-RMS noise) ratio for such a system is given by

$$S/N = \frac{(C_t)(\alpha_a)[\text{MTF}(L)][\text{MTF}(S)](G)}{\text{Noise}} \quad (3)$$

where C_t is the true contrast of the target with respect to background in the spectral region of the sensor; α^a is the atmospheric attenuation in the spectral region of the sensor; MTF(L) is the modulation transfer function of the lens at spatial frequency of interest; MTF(S) is the modulation transfer function of the sensor at spatial frequency of interest; and G is the sensor response to the signal present at the image plane.

The following yardstick can be observed from the above equation:

- If the equation results in a S/N <2, target may be detected but unable to track.
- If the equation results in a S/N >2, target may be tracked by both centroid and correlation tracker.
- If the equation results in a S/N >4, target may be tracked even by an edge tracker.

3.4 Update Rate

The target is tracked based on movement per field and as the target moving with high speed will have large movement between frames, it will require a large update rate. The performance of AVTS system for aerial target is determined by detection range of image system, scanning rate of stabilisation system, and track update rate. The maximum detection range for aerial target [target size: 2.5 m x 3 m, ΔT (temperature difference between target and ambient) = 20° , Visibility =15 km] for a typical TI (thermal imager) system with instantaneous field of view (IFOV) 0.1 mrad, is 14 km.

For tracking aerial target at longer ranges fast update rate is required. The target movement per field is calculated for different target speed is given in Table 2. For example target moving with 1 mach speed (300 m/s) will move by 6 m in one field period of 20 ms. For a sensor with 0.1 mrad IFOV the spatial resolution at 2 km range is 0.2 m, so the 6 m/field movement will require a update rate of 30 pixel/field. It can be seen that at shorter ranges, large track update rate is required which becomes the limiting factor for overall system performance. The maximum tracking rate for correlation tracker are of the order of ±4 pixels/field because of computational requirement[10].

At longer ranges due to small target size there is no feature available for correlation algorithm and at shorter range large update rate is required so correlation tracker is not suitable for tracking aerial targets. The centroid tracker can follow highly dynamic bounded targets. Even when servo platform cannot follow the target acceleration or unpredicted motion, centroid algorithm can track the target, which may results in target motion in the image. So centroid tracking is best suited for tracking aerial targets.

3.5 Thresholding Algorithm

A fixed threshold, which compares a pixel value to a constant, is not suitable for varying background and scene-to-scene target contrast variation. A variable thresholding scheme, in which the threshold varies with the average value of the scene is used in AVT. The continuous time varying video signal is transformed into a discrete format suitable for processing, using variable thresholding algorithm. The threshold level can be determined by

$$V_{th} = (V_{peak} - V_{ave})N + V_{ave} \quad (4)$$

where N is the can vary from 0.1 to 0.9; V_{ave} is the average video level in the gate; V_{peak} is the peak video level; *and* V_{th} is the threshold level.

Table 2. Track update requirement for aerial targets

Range	2 km	3 km	4 km	5 km	8 km	10 km	15 km	20 km
Target speed	**Target movement in 20 ms (pixels/field)**							
0.5 Mach	15	10	7.5	6	3.75	3	2	1.5
1.0 Mach	30	20	15.0	12	7.50	6	4	3.0
2.0 Mach	60	40	30.0	24	15.00	12	8	6.0

To determine V_{peak} and V_{ave}, several functions have to be taken into account[10]. First V_{ave} depends upon the area of the gate for a linear response with small and large gate size. Large gate has a faster response time than a smaller gate. However, it is advantageous to allow only small frame-to-frame variation in V_{th}. Therefore, the V_{th} for automatic thresholding is given by

$$V_{th} = [(V_{peak} - V_{ave})N + V_{ave}]e^{-t/\tau} \quad (5)$$

where $\tau = RC$. C has a constant value, R is a function of gate area, and t is the frame rate.

Figure 3 shows the variation of V_{th} with V_{ave} at different values of N with $V_{peak} = 150$ and $\tau = 0.10$. It is quite obvious from this figure that for adaptable threshold method, the selection of N is critical. If the difference between V_{peak} and V_{ave} is small, it is advisable to choose value of N to a higher side so that little variation in V_{ave} in successive frames should not affect the tracking using automatic thresholding.

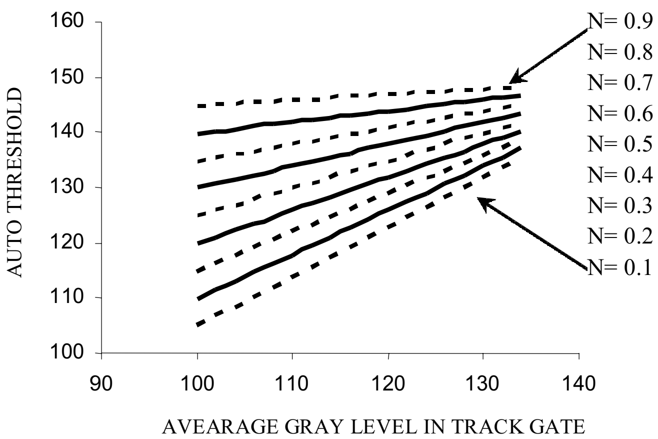

Figure 3. Variation of V_{th} with V_{ave} for different N.

4. TRACKER SYSTEM PERFORMANCE CRITERIA

The performance of a tracking system is determined by aiming point error, duration of successful tracking, number of tracking losses, and indication of confidence and system reaction time. Other performance criteria include the rate of convergence to steady-state tracking, robustness against countermeasures, and the ability to deal with multiple targets. The performance of AVT system is analysed based on these parameters.

4.1 Aiming-point Error

It measures how precisely the tracker follows a target. The absolute aiming point error is defined as difference between the perceived target positions as determined by the tracking system and the ideal aim point on the target. Error on the x- and y-axes are measured separately. The normalised aiming-point error defined as the absolute error divided by target size is generally used for comparing the tracker performance.

The instantaneous track error $e_t(k)$, represented by the algebraic sum of three error constituents

$$e_t(k) = e_b(k) + e_d(k) + e_j(k) \quad (6)$$

where $e_b(k)$ is the systematic bias term; $e_d(k)$ is the drift error representing error that grow with time; and $e_j(k)$ is the random pointing (jitter) error resulting from noise.

The bias term can arise from mechanical, servo, and/or gyro errors. Offsets in the target location estimation algorithms can also lead to bias error. Drift error may result from adaptation in the target location algorithm and/or from gyro drift. The jitter error arises from noise in the sensor data, round off errors in the processor, and sensor platform disturbances.

A stabilisation system is required to compensate for disturbances arise from the servomechanism, the electronics, the optics and the imaging part of the AVTS and brings the LOS jitter within tolerance limit[9]. The tolerable jitter limit may vary from few μrad to 100 μrad depending on overall FOV and other factors. The stabilisation can be done by active or passive means. The high frequency disturbance can be attenuated by passive means like shock absorber. To attenuate the low frequency LOS jitter, active stabilisation loop is required. The effective low frequency range of the active stabilisation system is determined by its bandwidth. As shown in Fig. 4, there are two stabilisation loops the rate loop and the pointing loop. Corresponding to the two loops the stabilisation system has two-loop bandwidth, the rate loop and pointing loop. The stabilisation loop is an inner loop to pointing loop and has the dynamic characteristics of the mechanical drive, motor and power electronics and thus determines

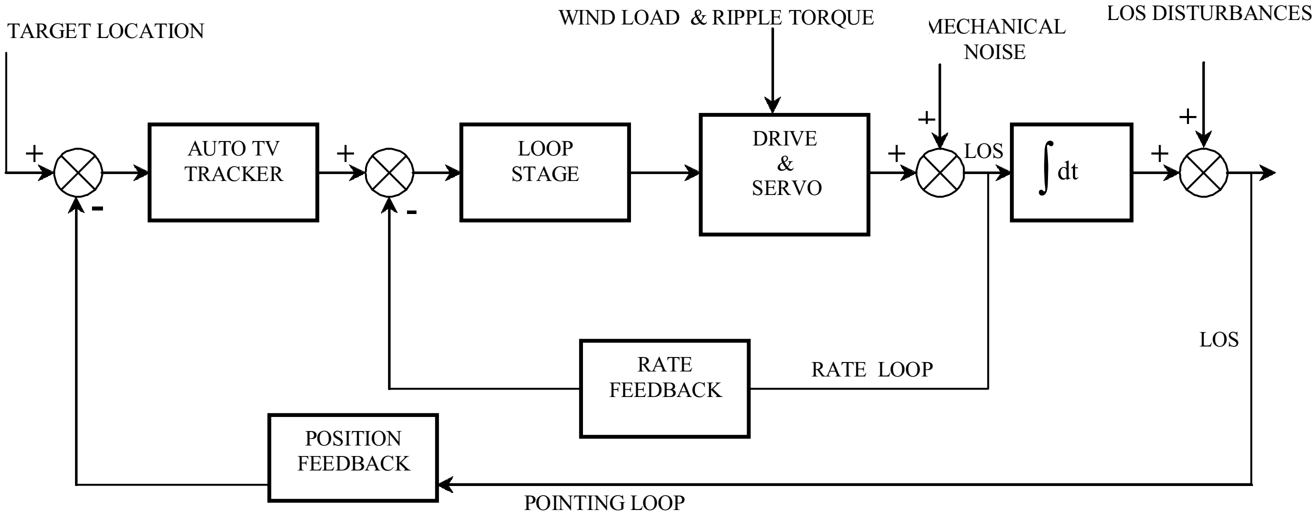

Figure 4. Target tracking control loops.

the dynamic behaviour of the AVT. To achieve a better AVT performance, one has to increase this bandwidth and achieve higher stabilisation loop gain. This enables to attenuate the external LOS disturbance to smaller value.

4.2 Duration of Successful Tracking

A target is being successfully tracked when the tracker system maintains a constant lock on a particular visual target and the aim point error does not exceed the limit. Tracking starts when the tracker detects automatically or is locked on to an object. Tracking ends when either the tracker decides that there is no longer an object to be tracked or when the radial aim point error exceeds the size limit given in the reference database. Duration of successful tracking (DOST) is defined as

$$\text{DOST} = \frac{\left(\begin{array}{c}\text{Number of frames in which a}\\ \text{target is successfully track}\end{array}\right)}{\left(\begin{array}{c}\text{Number of frames in which}\\ \text{target is visible}\end{array}\right)} \qquad (7)$$

The number of failures of tracking system during the possible tracking phase is defined as tracking loss. A failure can be a change of track to another target or background object or the result of the aim point error distance exceeding the tolerance limits. It is quantified by loss of lock probability (P_L) and can be caused by extreme target/platform motion and/or by imagery in which target's location cannot be reliably estimated. Loss of lock caused by target motion usually results from a poorly performing LOS determination function and/or the gimbals control system.

The specific cause can be found by the hybrid analysis method, which employs characterisation of target location estimator performance. Analytically it is difficult to determine the conditions for loss of lock resulting from an unreliable estimate of the target's location. The primary means of determining P_L, for this case is simulation of the image processing algorithms using imagery containing the particular scene factor to be investigated. Scene factors that may result in loss of lock include significant target signature changes, target obscuration by ground clutter, multiple targets in a crossing manoeuver, thermal merging or loss of contrast between parts of target and the background, etc.

4.3 Indication of Confidence

The reliability of the measured target position is very important. In case of low probability of pursuing the target, the system has to indicate a mode of prediction. At the time of prediction, the tracker estimates the target position by additional strategies to ensure target lock on. In correlation algorithm, depending on percentage of matching, a confidence level can be defined. In centroid mode of tracking, the confidence level can be defined in terms of target/ background contrast and growth of target size in successive frames.

Table 3. Target/background conditions versus tracking algorithms

Target background algorithm	Aircraft sky	Aircraft terrain	Ship sky/sea	Ship shore	Vehicle on land	Terrain feature	Missile	Re-entry vehicle	Star (point source)
Edge	0	-	X	0	-	-	X	X	-
Centroid	X	0	X	0	0	X	X	X	X
Correlation	0	X	X	X	X	X	X	X	0

X: typical application; 0: Possible application; -: Not a typical application

4.4 System Reaction Time

The delay time a tracker incurs in calculating the target position and delivering the result to the fire control system is defined as system reaction time. Normally processing is completed in one frame time after which the target error signal corresponding to new target position is fed to stabilisation system. To improve the confidence level and image contrast, various preprocessing techniques can be incorporated, which may introduce one frame throughput delay.

5. DISCUSSIONS

Various performance measures have been defined to evaluate the AVT system and performance of target-tracking algorithms, under various target engagement conditions has been analysed. The requirement of other sub-systems like sensor, stabilisation system, etc., to meet the overall system requirement has also been discussed. Based on above performance measure the selection criteria for tracking algorithms for different scenario is given in Table 3. The centroid algorithms have moderate computational load and it can follows the highly dynamic target. Centroid tracker always maintains the aim point at the centre of the target while in the correlation tracker, the aim point is defined at the window centre and vary with target manoeuverability. Therefore, centroid tracker is better suitable for fire control application where there is a requirement of preciously defining the aim point. It can be concluded from the above discussion that centroid tracking is more suitable for tracking fast moving targets like aircraft, missiles whereas correlation tracker has better performance for ground targets against terrain backgrounds. The edge tracker is best suited for tracking unbounded target and it can also track high dynamic target but it require more SNR.

ACKNOWLEDGEMENTS

The authors are grateful to Shri S.S. Sundaram, Director, Instruments Research and Development Eastablishment, Dehradun, for granting permission to publish this paper.

REFERENCES

1. Michael, Koligman. & Joseph, Dirbas. Automated video and infrared tracking technology. *SPIE,* 1999, **3808**, 44-51.
2. John, E. Albus; Lloyd J. Lewins & Julie, R. Schaht. Centroid tracking using probability map for target segmentation. *SPIE*, 2002, **4712**, 175-85.
3. Joseph, J. Fitchek; John, P. Lee & Herring, David. PICES: A multi-source tracker/correlator system. *SPIE*, 1989, **1096**, 258-68.
4. Downey, George. Electro-optical tracking systems considerations. *SPIE Proceeding*, 1989, **1111**, 70-84.
5. Mori, S.; Chang, K.C.; Chong, C.Y. & Dunn, K. P. Tracking performance evaluation-prediction of track purity. *SPIE*, 1989, **1096**, 215-23.
6. Sarma, C.G.S.; Venkatgiri Rao, S.; Badrinarayanan, M.S. & Mishra, R.K. Estimation of target acquisition and tracking system performance characteristics from aerial imagery. *SPIE*, 1994, **2221**, 335-44.
7. Dudzik, Michael, C. (Ed.). Electro-optical systems design, analysis and testing. The Infrared & Electro-optical Systems Handbook, Vol. 4, Chap. 5, 1993, .
8. Gerald, C. Holst. Electro-optical imaging system performance. SPIE Engineering Press, 2000.

9. Larry, A. Stockum & George, R. Carroll. Precision stabilised platform for shipboard electro-optical system. *SPIE*, 1984, **493**, 414-25.

10. Automatic Television Target Tracker: Reference Manual. DBA Inc, 1998.

11. Automatic TV Tracker ADEPT 30: User Manual. Octec Inc, 2004.

Contributors

Mr B.S. Chauhan received his MSc Physics (Electronics) from the Garhwal University, Srinagar (Garhwal), and MTech (Transducers Physics and Instrumentation) from Indian Institute of Technology (IIT) Madras, Chennai, in 1985 and 1994, respectively. He joined DRDO in 1985 and since then he has been working at the Instruments Research and Development Establishment (IRDE), Dehradun, in the area of electro-optical search, track and fire control system. His areas of research are real time image/signal processing and auto-target tracking system. He is a recipient of *Institute Merit Prize* from IIT Madras, for best academic record during MTech and *Laboratory Scientist of the Year Award* 2006.

Mr Manvendra Singh received his MTech (VLSI) from IIT Kanpur, and BTech from BIET, Jhansi. He joined DRDO in 1996. Since then he has been working at IRDE, Dehradun, in the area of electro-optical surveillance and video tracking system.

Mr V.K. Sharma received his AMIE (Electronic and Communication Eng.) from Institution of Engineers, Kolkata, in 1995. He joined DRDO in 1987. Since then he has been working at IRDE, Dehradun, in the area of electro-optical instrumentation.

Mr P.C. Pandey received his MSc (Mathematics) from the Garhwal University, Srinagar (Garhwal), and ME (Computer Science) from the Delhi University in 1987 and 2000, respectively. He joined DRDO in 1987. Since then he has been working at IRDE, Dehradun, in the area of electro-optical instrumentation.